

#7

RECEIVED

FEB 25 2002

TECH CENTER 1600/2900



1634

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/614,150

DATE: 02/12/2002

TIME: 13:03:36

Input Set : F:\SEQLIST 43006+2.TXT

Output Set: N:\CRF3\02122002\I614150.raw

ENTERED

```

4 <110> APPLICANT: Venter, J. Craig
5   et al.
7 <120> TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
8   ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
9   DROSOPHILA GENES.
11 <130> FILE REFERENCE: CL000728
C--> 13 <140> CURRENT APPLICATION NUMBER: US/09/614,150
C--> 13 <141> CURRENT FILING DATE: 2000-07-11
13 <150> PRIOR APPLICATION NUMBER: 60/157,832
14 <151> PRIOR FILING DATE: 1999-10-05
16 <150> PRIOR APPLICATION NUMBER: 60/160,191
17 <151> PRIOR FILING DATE: 1999-10-19
19 <150> PRIOR APPLICATION NUMBER: 60/161,932
20 <151> PRIOR FILING DATE: 1999-10-28
22 <150> PRIOR APPLICATION NUMBER: 60/164,769
23 <151> PRIOR FILING DATE: 1999-11-12
25 <150> PRIOR APPLICATION NUMBER: 60/173,383
26 <151> PRIOR FILING DATE: 1999-12-28
28 <150> PRIOR APPLICATION NUMBER: 60/175,693
29 <151> PRIOR FILING DATE: 2000-01-12
31 <150> PRIOR APPLICATION NUMBER: 60/184,831
32 <151> PRIOR FILING DATE: 2000-02-24
34 <150> PRIOR APPLICATION NUMBER: 60/191,637
35 <151> PRIOR FILING DATE: 2000-03-23
37 <160> NUMBER OF SEQ ID NOS: 43008
39 <170> SOFTWARE: FastSEQ for Windows Version 4.0
41 <210> SEQ ID NO: 1
42 <211> LENGTH: 2397
43 <212> TYPE: DNA
44 <213> ORGANISM: DROSOPHILA
46 <400> SEQUENCE: 1
47 aagtgaagcc aagtgaaaat gaaaacaaag tttaacaaac cgtggctaac gagctacatt 60
48 attaaaaactt caaaaatgca taaatttgcc gaagacgaat ggagcgaact gcagactgcg 120
49 gcattataat aatataatgc atatcctggg caggcatttg aaagttattg acacctcctc 180
50 ctctgctcac cgctctacct actcccatca cccaacagcc ggtaaaactc gtgacgtgat 240
51 cggtttttgag gcacctacga aggagagtgt tgcaacacag gccagggttc acaaaagctg 300
52 cacgatctgg gcactgggca cccactcccc aatccacaat cccaaccccc actctccatt 360
53 ccaatcctag cttcaatccc agcccaactg gaactgcgga tcgagttcgc gactcggctg 420
54 ccaagtgatt ctgcgcatat tagttggcca ttaggcgcac accattgtca gctgtgaacg 480
55 gcagtcgaaa gcgcattaaa attcagacgg tggttaatat tacaagcgcc ccctgttgca 540
56 gaaaaggggg aaagggtttcg cgctgatttg cagctgccac tcaagagggt ttttggaggg 600
57 ggtggtgctg cagatcgctg aggagcttgt gaaacgtatt ggcgacgtga tatgttaac 660
58 tttcaggcga aaacgtggag cgttttggat gcggatgttt ggcagacaga aaagcacttg 720

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/614,150

DATE: 02/12/2002

TIME: 13:03:36

Input Set : F:\SEQLIST 43006+2.TXT

Output Set: N:\CRF3\02122002\I614150.raw

```

59 tcggttagac atggttgga atgaagattc ctcccttttt gatccaatga atgcaatata 780
60 atactttttt tgggtttttac gaaaagcgtt cacttatctt agcacttcgc tttcattgca 840
61 gactggactc actaaatatg tatcttagct ttataggaaa catttttccg caatagttta 900
62 tgctcttcaa aggcctttcag ttggtaatag caatcgcca actgaaaact ttatttaata 960
63 cttgtgtggc aattataaac agcataatat gtgtctaaatg ctaacccttt atgttaagaa 1020
64 aagtatcttt ggcatcttga accaccaaaa cggatctctt ggcatctggt gtaatgagca 1080
65 cctcgtgata cttgggtacga actctcagta gagtcaaatt ttgggctggc tccatttttg 1140
66 agaggaacgc ctggcacttc tcccgcaaat tgtcatatag acaggcgtac tgcaagccct 1200
67 cctgacgac catcgagggt ttcaccggca cacccgagt attcatgac aatatgtcct 1260
68 ccacgccggg tttctcctgc actttgcgaa atacttcttc cacgtaactt ttagttcgct 1320
69 ttggctgcgt aaagttaaac tatcagttag gatgtaaata catttaacat acttacatct 1380
70 gtctctgctt cctgcatttt gcctacagtt ttttttaaac gtccaccaa tgtactttta 1440
71 gttataaatt ttctaagtca gtaactgctt ggggttgctta aatttgagtt tatgggtatg 1500
72 acattaatgt attggagtac cttgaattga ttgtagttag actcgagtaa aatcgagta 1560
73 attgcgcacc gttattagta tccagccaga agcaaagttt aataaattcc ttgggtatca 1620
74 tcctgtgacg tcataatcac gggcattaga tgatttagtc ggatgccaa cagggccaaa 1680
75 ttaggcggga gttatttgcc aaagcgggct gtaagtctat ggccaactct aatcgctctc 1740
76 taactcagct gccaaagcat atggacggcc cactttcgag gcgcgtccc agagcataaa 1800
77 gtgcaattat agcgggtgag cgaggagcag ggccatgaaa taccaacacc cgacagctgt 1860
78 aaacaacaat ttgccgggag gccggccata agtgcatctc gcatctgaga ccaaatgcca 1920
79 ataactcgcc gtgcgagagc gcttataaaa aagtatttta aaagcccagc caactcgggt 1980
80 ggagtccaag agccaaactc aagccaatct caagccaggc caaaccaagc caatccaagc 2040
81 caacaatggc aaatgatgat gagacgtcca gtcgtttcac gttaaacgct aagcgtttcg 2100
82 gcaaacgcgc aaaaagttag acgggtgccaa ggagccgact tcactcaaga cgagttgagg 2160
83 atggaaaactt gaatgaaaat acctcgaaaca ggagaagcga acgtaatgag caaatttatg 2220
84 gtcaattgat aatttagagg cattcggtat catcaataat ggacccaacc gctgtggctc 2280
85 tggatcgaag atgaggcatt gttattgtta ccaatgtggc gcaaattgtg ctaaattgtt 2340
86 cggcagaccc tacgcactcg cacatgaaat atatttatgc gagcatctta aaggggc 2397

```

88 <210> SEQ ID NO: 2

89 <211> LENGTH: 345

90 <212> TYPE: DNA

91 <213> ORGANISM: DROSOPHILA

93 <400> SEQUENCE: 2

```

94 atgcaggaag cagagacaga tccaaagcga actaaaagtt acgtggaaga agtatttcgc 60
95 aaagtgcagg agaaaacccg cgtggaggac atattgatca tgaatcactc ggggtgtgccg 120
96 gtgaaaacct cgatggatcg tcaggagggc ttgcagtacg cctgtctata tgacaatttg 180
97 cgggagaagt gccaggcgtt cctctccaaa atggagccag cccaaaattt gactctactg 240
98 agagttcgta ccaagtatca cgagggtgctc attacaccag atgccaaagat caccgttttg 300
99 gtgggttcaga atgccaaaga tacttttctt aacataaagg gttag 345

```

101 <210> SEQ ID NO: 3

102 <211> LENGTH: 114

103 <212> TYPE: PRT

104 <213> ORGANISM: DROSOPHILA

106 <400> SEQUENCE: 3

```

107 Met Gln Glu Ala Glu Thr Asp Pro Lys Arg Thr Lys Ser Tyr Val Glu
108 1 5 10 15
109 Glu Val Phe Arg Lys Val Gln Glu Lys Pro Gly Val Glu Asp Ile Leu
110 20 25 30
111 Ile Met Asn His Ser Gly Val Pro Val Lys Thr Ser Met Asp Arg Gln

```

RAW SEQUENCE LISTING

DATE: 02/12/2002

PATENT APPLICATION: US/09/614,150

TIME: 13:03:36

Input Set : F:\SEQLIST 43006+2.TXT

Output Set: N:\CRF3\02122002\I614150.raw

```

112          35          40          45
113 Glu Gly Leu Gln Tyr Ala Cys Leu Tyr Asp Asn Leu Arg Glu Lys Cys
114          50          55          60
115 Gln Ala Phe Leu Ser Lys Met Glu Pro Ala Gln Asn Leu Thr Leu Leu
116 65          70          75          80
117 Arg Val Arg Thr Lys Tyr His Glu Val Leu Ile Thr Pro Asp Ala Lys
118          85          90          95
119 Ile Thr Val Leu Val Val Gln Asn Ala Lys Asp Thr Phe Leu Asn Ile
120          100          105          110
121 Lys Gly
125 <210> SEQ ID NO: 4
126 <211> LENGTH: 4064
127 <212> TYPE: DNA
128 <213> ORGANISM: DROSOPHILA
130 <400> SEQUENCE: 4
131 ccaccgatgc cagctcgcag atcgcagctc ccaggacccc acttggcgca agttttcgga 60
132 cccgctgcac tcctcgatgg cggcactgag ctccagttgg gtcatatgct catttgtgtg 120
133 catgacgcac tgacctgttt ctaatgcaac aaaaacgtta agcgcctgcc agccactcgc 180
134 agtgcagtca actgtgtcct cttggctcag agagcacagc caccacacca cttgagcgcc 240
135 ccaaatccca tacaccgcta taccagctc caccacactt ggccatgttc tttggtgcct 300
136 cgcagccgta gcatactttt ttacgtgtgc tgccggagag tttattttac gcacgccttg 360
137 accagacgac ggaggcggat gcagccttcc actttgcact ccgcaggcaa gtggaaaagt 420
138 caagcggata aacaagagat taaaatgtca taaacaatgc tttttgccac tccggccaag 480
139 caacttaatg aaagccggga aaatgggttt cctccgctgc agctgcaaag tttcaacttc 540
140 aagtttggat ctgaaagcgg agtatctctc gaaactcctg tcgatgcaac tcaataaatt 600
141 tccattaaac ttttgatttg tttccatgac tttttccgcc ctgcaagacc agagaattca 660
142 attgaaattg tttccaaaac aagctacaca aaaagttggt aacaactatt tgaatgccag 720
143 ccaataaatg ggtcttttgc ctatgtgact atattgtaat tcctaataca aatcggagtc 780
144 ggagagacaa taataaggct atatgcacat actttatctt aagctgaaaa acttgtgaat 840
145 tattacatcg aaatcttcaa aagattgatc aagcagtatg ttccaaatat tttttattat 900
146 cgttatagtt acgtttccat tattacgaac atttaaaata tttatttgca ggtaagggca 960
147 tagaaccatc gacttctcga aagtaaaaac ttttcatttg taacacaatt gaaaaaaact 1020
148 ttattttcgt tgggcctgcg aacaacagga ggggtgcaca aatgacaaac tatcatacag 1080
149 aggcataaat tagtctcgca tatttgcgcc aaacaagaaa ctataaaaca cttagatata 1140
150 cttcggatta gatttactgc ccatgtgcac acacacacac gcgcgcgctc ttaaacacac 1200
151 acacacatgt gcgtctttta aagtcccttc gataaacgac atttaggggc tattctaaat 1260
152 gaaatactgg atgctgcgtc ggtgcaagca aaacctcagg caatttgctg gcgacaagaa 1320
153 tcagcaacgc tgtgatgctg aacgtaaaact gatccttggc ggcgaaaatt acggatacga 1380
154 tacaattgtg gaggactaat ttcccgttta tgtgttatgt gtaatgcgtg tacggtatta 1440
155 agtattgatt cgttaacata tacatgtgga tagactgaag gaataaagga atgaatattg 1500
156 agtgcgcatt gaacaatcaa ataaggtgtg aaccgatgat gtaggcggtt tactcggcgg 1560
157 cggcattgat ggtgatagtc ttgcccttgg gatcgtcgaa gcggtccatc gtgttgatgt 1620
158 ccatcatcca gcagacgccg acgaacatga tgaccagcag tagggcgacc acaaacagtc 1680
159 ccatcaggat accgggcgtc acgaagccca cacagtccca ggagtcgccg aagggaaagt 1740
160 cctccttgta cgtaccatcg aagggagcct gcagctgcag ggacttgaag ctcaggggtg 1800
161 tgtacatggt gttgacggcc gcggactcga gggtgagggt gccgcacgag taagagaacg 1860
162 tggttggggc attcacgccg ctggagcgga agttgttggt gtcgtagacc aaattgctca 1920
163 gactaaagta tccgccatta taaacaacgt caaaggatgat tggcttgctg gcgacagatg 1980
164 tgtccataac aacggatagc ttggtagagc tagagttggt gagcttgagg tcggtgacgg 2040

```

RAW SEQUENCE LISTING

DATE: 02/12/2002

PATENT APPLICATION: US/09/614,150

TIME: 13:03:36

Input Set : F:\SEQLIST 43006+2.TXT

Output Set: N:\CRF3\02122002\I614150.raw

```

165 tgatgggggtt gccgtttag agcagggcag tgtagaagat ctgaaactga ttggtcgact 2100
166 tccacatgat gccaccggtt gtggcggcgg tgtcacggcg ggtacgacgc tgcaccacag 2160
167 gagccgtgga gggggcggcc aggtacagat aagccacctt gcactcgaac tgcttgctga 2220
168 tagcagctat ggcggcgtct aaatgtgtta aatcattggt aaattagtgt gtcttattga 2280
169 cgcctttgta gtacgttatt gttgaccac catgtgactc cagactggcc tcccggctct 2340
170 cggcggcgtc ctcgaaggct acaaacaggg ccgtacctac ggcacacttg gccggcgtgg 2400
171 tcagctttcc actggcatcg atggagtgtt gctcgcgctt ggcggccact gagcgcagt 2460
172 cctccgaggg gttctccacg ctggtgtagt aggtcttggg gctcactccc tgcagctggg 2520
173 cgtagcagga ctgcgcctgg gagttggagc acaggaagtc cttgctgctc agctgagttg 2580
174 gggaacaaac acatgtgtca caagccctag ctatctaagt gggttctcaa gcacttacgc 2640
175 cattttcctc gaaggccacg accatgtgat cttccagcaa tgcagccaac tgctcggcaa 2700
176 actccacttg ggacaccgtc ttcagggagg gtttcgccac actgtgaaaa aaggaggag 2760
177 gcactcgta cactcacatt caactggtgc aggcgtcac ccacctgtt gctccccaca 2820
178 aaaagacggg cgtttgctcc gccacggcag ccccaatgac gcacaacgca atcagcgatt 2880
179 tccacaacat cgcgctggat tcgcttactt cggtgggggg ttaatggaaa caaaggagaa 2940
180 cgaaaaatgc ataagtaagg aaccaaggag aacgcaagag gcactaaaaa ccaaccgaaa 3000
181 agtgaactga aactacaaac gaaaaatggt ggtgaaaatt ttcgctttgc cttgctcgac 3060
182 gtgccttatt ccagtggaga aaaaaacctt atgacgggtc tgtgatgcgc accagctgac 3120
183 cgcagtgtcg tctaaaaata ccaaaagcgt gagattagga aaatactaga aaaaaattat 3180
184 tcgactctga gagcatttag taaagcctaa ttggatattt gcaccaggaa ctgacattaa 3240
185 ctgggtttct tttcaaaatt aatacttata agaaagtata gtttgtcaaa taaggttcac 3300
186 atcttcggtg tccaattagc gcagtcacac acctttccag ctggaaactg tcaacggaac 3360
187 cggttccggt aattccacat gttgattctt taaaaccgct atgttctatt tcgcactgct 3420
188 tactaatagt ttaattaaat atctaaagaa attcgggttt tatactgatt gcttttataa 3480
189 gctcgatttg aaccgagttt ttatttcagt ctttccagtt agtaatttta cgcagttcta 3540
190 gcacggccac actgtcccgc agggaaaaca tgtttgttat gattttactg cgggaattgt 3600
191 gtgtgaaatg tgccatttga ggacctgaga atgcgattac gccgcggtg gccgcgaagg 3660
192 atgcggcgtg ccatggaaca actgcgtctc cagtcgccgc ggaagctgct ggcgctcacg 3720
193 agtgagttac cttgctcacc agccccatta cttaactaat gttctttttt tttcaatttt 3780
194 agtgactgga ggactctgcc tggctatctg ggtagttgcc aactatctgt ccgcatcgga 3840
195 ggagggccaa accggcggct tcagcagcgc ctgcactgcc atcgatgcgg tttacacgtg 3900
196 ggtcaatggc tcggtatcca actttattga ggacatccgg cgattcgatg acaagtacga 3960
197 tccctcgcgg ttcgatgaca agaatgagct gcggtactct ctgaggtccc tggaaaaaca 4020
198 cgcgcgatgg atcaggcatg tgtacatagt aaccaatggc caga 4064

```

200 <210> SEQ ID NO: 5

201 <211> LENGTH: 1863

202 <212> TYPE: DNA

203 <213> ORGANISM: DROSOPHILA

205 <400> SEQUENCE: 5

```

206 cgacgtcgag caaggcaaag cgaaaatttt caccaccatt tttcgtttgt agtttcagtt 60
207 cacttttcgg ttggttttta gtgcctcttg cgttctcctt ggttccttac ttatgcattt 120
208 ttcgtttctc tttgtttcca ttaaaccctc accgaagtaa gcgaatccag cgcgatgttg 180
209 tggaaatcgc tgattgcgtt gtgcgtcatt ggggctgccg tggcggagca aacgcccgct 240
210 tttttgtggg gagccaacag tgtggcgaaa cctccctga agacgggtgtc ccaagtggag 300
211 tttgccgagc agttggctgc attgctggaa gatcacatgg tcgtggcctt cgaggaaaat 360
212 ggcctgagca gcaaggactt cctgtgctcc aactcccagg cgcagtcctg ctacgccag 420
213 ctgcagggag tgagcccaa gacctactac accagcgtgg agaaccctc ggaggcactg 480
214 cgtcagtggt ccgccaagcg cgagcacaac tccatcgatg ccagtggaaa gctgaccacg 540
215 ccggccaagt gtgccgtagg tacggccctg tttgtgacct tcgaggacgc cgccgagagc 600

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/614,150

DATE: 02/12/2002

TIME: 13:03:36

Input Set : F:\SEQLIST 43006+2.TXT

Output Set: N:\CRF3\02122002\I614150.raw

```

216 cgggaggcca gtctggagtc acatgacgcc gccatagctg ctatcagcaa gcagttcgag 660
217 tgcaagggtgg cttatctgta cctggccgcc ccctccacgg ctccgtgtgg gcagcgtcgt 720
218 acccgccgtg acaccgccgc cacaaccggt ggcatcatgt ggaagtcgac caatcagttt 780
219 cagatcttct acactgccct gctctacaac ggcaacccca tcaccgtcac cgacctcaag 840
220 ctcaccaact ctagctctac caagctatcc gttgttatgg acacatctgt cgccgacaag 900
221 ccaatcacct ttgacgttgt ttataatggc ggatacttta gtctgagcaa tttggtctac 960
222 gacaacaaca acttccgctc cagcggcggtg aatgccccaa ccacgttctc ttactcgtgc 1020
223 ggcaacctca ccctcgagtc cgcggccgctc aacaacatgt acaacaccct gagcttcaag 1080
224 tccctgcagc tgcaggctcc cttcgatggg acgtacaagg aggactttcc cttcggcgac 1140
225 tcctgggact gtgtgggctt cgtgacgccc ggtatcctga tgggactgtt tgtggtcgcc 1200
226 ctactgctgg tcatcatgtt cgtcggcgtc tgctggatga tggacatcaa cacgatggac 1260
227 cgcttcgacg atcccaaggg caagactatc accatcaatg ccgcccgcga gtaaacgcc 1320
228 tacatcatcg gttcacacct tatttgattg ttcaatgcgc actcaatatt cattccttta 1380
229 ttccttcagt ctatccacat gtatatgta acgaatcaat acttaatacc gtacacgcat 1440
230 tacacataac acataaccgg gaaattagtc ctcacaattt gtatcgtatc cgtaattttc 1500
231 gccgccaaagg atcagtttac gttcagcatc acagcgttgc tgattcttgt cgccgacaaa 1560
232 ttgcctgagg ttttgcttgc accgacgcag catccagtat ttcatttaga atagccccta 1620
233 aatgtcgttt atcgaaggaa ctttaaaaga cgcacatgtg tgtgtgtgtt taagagcgcg 1680
234 cgcgtgtgtg tgtgtgcaca tgggcagtaa atctaattccg aagtatatct aagtgtttta 1740
235 tagtttcttg tttggcgcaa atatgcgaga ctaatttatg cctctgtatg atagtgtgtc 1800
236 atttgtgcaa ccctcctgtt gttcgcaggc ccaacgaaaa taaagttttt ttcaattgtg 1860
237 tta

```

239 <210> SEQ ID NO: 6

240 <211> LENGTH: 379

241 <212> TYPE: PRT

242 <213> ORGANISM: DROSOPHILA

244 <400> SEQUENCE: 6

```

245 Met Leu Trp Lys Ser Leu Ile Ala Leu Cys Val Ile Gly Ala Ala Val
246 1 5 10 15
247 Ala Glu Gln Thr Pro Val Phe Leu Trp Gly Ala Asn Ser Val Ala Lys
248 20 25 30
249 Pro Ser Leu Lys Thr Val Ser Gln Val Glu Phe Ala Glu Gln Leu Ala
250 35 40 45
251 Ala Leu Leu Glu Asp His Met Val Val Ala Phe Glu Glu Asn Gly Leu
252 50 55 60
253 Ser Ser Lys Asp Phe Leu Cys Ser Asn Ser Gln Ala Gln Ser Cys Tyr
254 65 70 75 80
255 Ala Gln Leu Gln Gly Val Ser Pro Lys Thr Tyr Tyr Thr Ser Val Glu
256 85 90 95
257 Asn Pro Ser Glu Ala Leu Arg Ser Val Ala Ala Lys Arg Glu His Asn
258 100 105 110
259 Ser Ile Asp Ala Ser Gly Lys Leu Thr Thr Pro Ala Lys Cys Ala Val
260 115 120 125
261 Gly Thr Ala Leu Phe Val Thr Phe Glu Asp Ala Ala Glu Ser Arg Glu
262 130 135 140
263 Ala Ser Leu Glu Ser His Asp Ala Ala Ile Ala Ala Ile Ser Lys Gln
264 145 150 155 160
265 Phe Glu Cys Lys Val Ala Tyr Leu Tyr Leu Ala Ala Pro Ser Thr Ala
266 165 170 175

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/614,150

DATE: 02/12/2002

TIME: 13:03:37

Input Set : F:\SEQLIST 43006+2.TXT

Output Set: N:\CRF3\02122002\I614150.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application No
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date